Case Studies Synthesis: A Thematic, Cross-Case, and Narrative Synthesis Worked Example

Daniela S Cruzes [a, b], Tore Dybå [b], Per Runeson [c] and Martin Höst [c]

[a] *Department of Computer and Information Science (IDI), NTNU, Trondheim, Norway, NO-7491.*
[b] *Department of Information and Communication Technology (ICT), SINTEF, Trondheim, Norway, NO-7465.*
[c] *Department Computer Science, Lund University, Box 118, SE-221 00 Lund, Sweden.*
danielac@sintef.no, tore.dyba@sintef.no, per.runeson@cs.lth.se, martin.host@cs.lth.se

Abstract—**Case studies are largely used for investigating software engineering practices. They are characterized by their flexible nature, multiple forms of data collection, and are mostly informed by qualitative data. Synthesis of case studies is necessary to build a body of knowledge from individual cases. There are many methods for such synthesis, but they are yet not well explored in software engineering. The objective of this research is to demonstrate the similarities and differences of the results and conclusions when applying three different methods of synthesis, and to discuss the challenges of synthesizing evidence from reported case studies in SE. We describe a worked example of three such methods where three independent teams synthesized two studies that investigated critical factors of trust in outsourced projects through thematic synthesis and cross-case analysis, and compared these to each other and also to an already published narrative synthesis. In addition, despite that the primary studies were well presented for synthesis, we identified challenges in the use of case studies synthesis methods related to the goals and research questions of the synthesis, the types and number of case studies, variations in context, limited access to raw data, and quality of the case studies.**

Keywords—**Evidence-based software engineering; systematic reviews; research synthesis; case study**

I. INTRODUCTION

Software engineering (SE) projects, processes, and artifacts are typical objects for which case studies are a feasible research approach. Case studies are characterized by their flexible nature, evolving over the course of the study, focusing on a phenomenon in context, using multiple methods of evidence or data collection. Selection of cases to study is not governed by sampling logic and representativeness; rather cases are selected for the purpose of being 'typical', 'critical', 'revelatory', or 'unique' in some respect [32]. Case studies, as any empirical research is costly and it is usually not possible to investigate all the aspects of a phenomenon in one case study. The issues of what kind of generalization is possible from a single case and how such generalizations might be established are important to investigate, as these issues are not concerned with statistical generalization, where there is established theory and practice on how to generalize.

Progress, however, in any scientific field depends on the accumulation of knowledge from diverse aspects of a phenomenon; it is necessary, therefore, to adopt approaches for integrating and providing new interpretive explanations about existing case studies. Case study synthesis can help accomplish this goal, by extending the investigator's expertise beyond the single case [28] [29].

Research synthesis is used as a collective term for a family of methods to summarize, integrate, combine, and compare the findings of different studies on a specific topic or research question [4]. It is built upon the observation, that no matter how well designed and executed, empirical findings from single studies are limited in the extent to which they may be generalized [4]. The synthesis of case studies must take into account the flexible nature of the cases, the mixed qualitative and quantitative characteristic of

the data, and the type of cases being studied. The flexibility in the choice of methods for performing a case study is one of the characteristics that lead to challenges in conducting the synthesis.

The process of synthesis entails organizing the relevant evidence extracted from the included sources and then finding some way of bringing it together. The way the evidence is organized depends to some extent on the type(s) and scope of the evidence, the method(s) employed and on the preferences of the researcher [26]. As with data extraction, the process of organizing the studies is often facilitated by the use of charts or tables summarizing key aspects of the studies. The formats of these largely depend on how many studies or pieces of evidence are included, but they need to be capable of allowing repeated examination and comparison of the relevant data from each study.

Synthesis methods are usually tailored to a particular type of evidence, for example meta-analysis aggregates and averages different findings in experimental or quasi-experimental studies, whereas meta-ethnography synthesizes findings from qualitative studies [3]. In addition, there are a large variety of methods for synthesizing qualitative and mixed-methods evidence [4] [6] [26]. Common to these methods is that they embody the idea of making a new whole out of the parts to provide novel concepts and higher-order interpretations, novel explanatory frameworks, an argument, new or enhanced theories, or new conclusions. Further, many similar methods appear under different names in different research traditions. Cruzes and Dybå describe how some of these methods have been used in systematic literature reviews in SE [4], but the vast majority of the methods are yet unexplored in SE.

For the purpose of this paper three of the most relevant methods of case study synthesis are compared: thematic synthesis, cross-case analysis, and narrative synthesis. Our aim is to demonstrate the similarities and differences of the results and conclusions when applying different methods of synthesis, and to discuss the challenges of synthesizing evidence from reported case studies in SE. Our main research questions are:

> *What are the differences in the results when using narrative, cross-case or thematic synthesis of case studies evidence in SE?*

> *What are the main challenges of performing case studies synthesis in SE?*

To investigate these research questions, we performed two independent syntheses of two published case studies (on trust in outsourcing) [2] [24]. The primary studies were selected because of their relative homogeneity, allowing us to address the easier synthesis issues first. One team applied cross-case analysis of the two papers and the other team applied thematic synthesis. We compare and discuss the results of these two syntheses to each other and also to a third, already published narrative synthesis of the same two papers (Babar et al. [2]). In addition, we discuss the challenges of performing the syntheses. Preliminary findings were reported as a short paper at ESEM 2011 [7]. We have now explored the analysis in depth and present a worked example to illustrate the methods, and the challenges in applying them to published case studies.

The rest of this paper is organized as follows: Based on the literature on research synthesis we discuss case study synthesis and describe the three methods of synthesis in Section II. The worked example is described in Section III. The experiences, strengths and differences from the syntheses are presented in Section IV. Section V concludes and outlines further work.

II. CASE STUDY SYNTHESIS

Most case studies in SE research are single-case or few-case studies, with large sample comparative studies still being seldom. The result is that knowledge about the phenomena of SE practices, methods, and techniques are spread over a myriad of diverse studies. Additionally, the majority of the data collected in these case studies are observations and interviews that are analyzed qualitatively.

The simplest and possibly the most widely used way to combine such studies is the traditional informal, narrative literature review, which is used to review every kind of conceptual and empirical

literature, including case studies as well as quantitative studies. Relying primarily on the subjective insight and knowledge of the researcher, these traditional reviews lend themselves mainly to exploratory studies aimed at summarizing a certain research literature without applying a strict research question [26]. The advantage is that the researcher can put his/her own judgments of particular studies and compare them in a flexible manner. The disadvantage is that the researcher can be biased towards his/her own experience and beliefs on the topic. Besides, as traditional reviews typically do not develop clear criteria as to which studies are to be included and how they are synthesized, other researchers can hardly replicate their synthesis.

Systematic literature reviews (SLR) has been the approach used in SE for synthesizing research for diverse primary studies since 2005 [13][14]. In SLRs, the researchers attempt to gather relevant studies, critically appraise them, and come to judgments about what works using explicit, transparent, state-of-the-art methods. SLRs include details about each stage of the review process, including the questions guiding the review, search methods, inclusion and exclusion criteria, details on the data extraction and methods and process of synthesis. Synthesis is one of the phases in software engineering SLRs that suffer the most from lack of transparency and usage of state-of-the-art methods. Despite the fact that methods of synthesis have been available for many years in other disciplines [26], about half of the SLRs in SE limit themselves to map the area of study without synthesizing the evidence [4], and even the ones that do synthesize evidence are not fully exploring the methods that are well established in other disciplines.

For case studies in particular, synthesis methods have been available for at least four decades [16][17][21]. These methods allow systematic and rigorous synthesis of previous case-based research by generating findings and conclusions based on rich case material created by different researchers, contexts and study designs, and at the same time allowing for a much wider generalization than from single cases. The empirical evidence, which such syntheses depend upon, is the data on which a conclusion or judgment may be based. Although there are many ways to generate evidence, case studies have a special ability to provide deep understandings of the phenomena under study from direct observations of practice through rich, longitudinal and multi-sourced data. The synthesis must take into account the flexible nature of the case study, the qualitative and mixed characteristic of the data, and the number and type of cases in each primary study.

Table 1 outlines some of the methods that are most relevant for synthesizing evidence across case studies (a more complete list is provided in Cruzes and Dybå [4] [6]). Largely depending on the research goal and overall research approach, for the synthesis of qualitative case studies, most probably no single method will offer all the required features for performing the synthesis, so a combination of methods may often be the best approach. In the following, we describe and compare three most used of such methods; thematic synthesis, cross-case analysis, and narrative synthesis, which we use in the worked example to explore some of the methodological challenges of SE case studies synthesis (see Table 2).

Thematic synthesis is a method for identifying, analyzing, and reporting patterns (themes) within data. It is one of the most common methods for synthesis of evidence in SE [5]. Thematic synthesis resembles some of the characteristics of grounded theory analysis, in that the themes emerge from (are grounded in) the primary data. It minimally organizes and describes the data set in rich detail and frequently interprets various aspects of the research topic. It comprises the identification of the main, recurrent or most important (based on the specific question being answered or the theoretical position of the reviewer) issues or themes arising from a body of evidence [5]. The level of sophistication achieved by this method can vary; ranging from simple description of all the themes identified, through to analyses of how the different themes relate to one another in a conceptual map [26]. The advantage of thematic synthesis is that it provides a means of organizing and combining the findings from a large, diverse body of research [26]. It can handle qualitative and quantitative findings, and it can be a deductive, theoretically driven approach or an inductive one, in which themes 'emerge' from the process of synthesis. However, transparency is usually criticized in thematic synthesis, since there are many

different ways to perform it. Recently, Cruzes and Dybå [5] extended existing approaches of thematic synthesis with relevant guides and recommendations, conceptualize thematic synthesis in SE as a scientific inquiry consisting of five steps based on the extent literature (See also Table 2).

TABLE 1. RELEVANT CASE STUDY SYNTHESIS METHODS (ADAPTED FROM [4] [6]).

Synthesis method	Description	Strengths	Challenges
Case survey [16][17]	Formal process for systematically coding relevant data from a large number of case studies for quantitative analysis, allowing statistical comparisons across studies. Study findings and attributes are extracted using closed-form questions for increased reliability, while survey analysis methods are used on the extracted data. The resulting dataset is used to construct cross-case matrices or summary tables.	• Can incorporate diverse evidence types. • Can cope with large numbers of primary studies. • Could be used for theory-building.	• Applicable to outcomes, but less adequate for process. • Lacks sensitivity to interpretive aspects of evidence
Qualitative comparative analysis (QCA) [27]	The qualitative comparative analysis method is a mixed synthesis method that analyzes complex causal connections using Boolean logic to explain pathways to a particular outcome based on a truth table. The Boolean analysis of necessary and sufficient conditions for particular outcomes is based on the presence/absence of independent variables and outcomes in each primary study	• Transparent. • Can incorporate diverse forms of evidence. • Allows competing explanations to be explored and retained and permits theories about causality. • Does not require as many cases as the case survey method.	• Focused on causality determination, not interpretive aspects of qualitative data.
Cross-case analysis [19][20]	Includes a variety of devices, such as tabular displays and graphs, to manage and present qualitative data. It includes meta-matrices for partitioning and clustering data in various ways. Evidence from each primary study is summarized and coded under broad thematic headings, and then summarized within themes across studies with a brief citation of primary evidence. Commonalities and differences between the studies are noted.	• Highly systematic method. • Potentially allows inclusion of diverse evidence types. • Could be used for theory-building.	• Can be seen as unnecessarily and inappropriately stifling interpretive processes.
Thematic Synthesis [5][31]	A method for identifying, analyzing, and reporting patterns (themes) within data. It organizes and describes the data set in rich detail and interprets various aspects of the research topic. It can be used within different theoretical frameworks, and it can be an essentialist or realist method that reports experience, meanings, and the reality of participants. It can also be a constructionist method, which examines the ways in which events, realities, meanings, experience, and other aspects affect the range of discourses.	• Flexible procedures for reviewers. • Copes well with diverse evidence types. • Could be used for theory-building.	• Lack of transparency. • Largely descriptive/data-driven basis to groupings.
Narrative synthesis [25]	A defining characteristic of narrative synthesis is the adoption of a narrative (as opposed to statistical) summary of the findings of studies. It is a general framework of selected narrative descriptions and ordering of primary evidence with commentary and interpretation, combined with specific tools and techniques that help to increase transparency and trustworthiness. It can be applied to reviews of quantitative or qualitative research as individual tools and techniques can be selected according to the type of study design and data included in the review.	• Can cope with large evidence base, comprising diverse evidence types. • Flexibility. • Can be used for theory-building.	• Lack of transparency. • Many variants and lack of procedures/standards. • May be dependent on prejudices of reviewer.

Cross-case analysis is a method that facilitates the comparison of commonalities and differences in the events, activities, and processes; the units of analyses in case studies. The term cross-case analysis is sometimes used as a general umbrella term for the analysis of two or more case studies to produce a synthesized outcome [12]. In some contexts, it has narrower meaning, referring to a specific method for performing the analysis, organizing the data from the cases in tables and graphs. We use the term in the specific sense, referring to a method to synthesize the findings of two or more case studies. Although

there are several cross-case analysis approaches and techniques available to the case study researcher [12], currently, cross-case analysis has not been applied in SE systematic reviews [4]. The cross-case analysis, as proposed by Miles and Huberman [19][20] is originally presented as a method to synthesize evidence from multiple cases *within* a multi-case setting, rather than a secondary analysis of *different* case studies. However, there is nothing in the method as such, preventing it from being applied in secondary studies. The drawback in the secondary study context is that the access to raw data from the primary studies is limited by the publication format; but nevertheless, a limitation common for all synthesis methods. Miles and Huberman's process [19][20] consists of three concurrent flows of activities: *data reduction, data display* and *conclusion drawing/verification* (see Table 2).

TABLE 2 - DETAILED DESCRIPTION OF THEMATIC, CROSS-CASE AND NARRATIVE METHODS OF SYNTHESIS

Thematic Synthesis	*Cross-Case Analysis*	*Narrative Synthesis*
Purpose: Progressive theming to form a chain of reasoning.	**Purpose:** Progressive tabling to form a chain of reasoning.	**Purpose:** Progressive linking to form a chain of reasoning.
Data Sources: Findings and interpretations of existing studies and relevant theory.	**Data Sources:** Findings and interpretations of existing studies and relevant theory.	**Data Sources:** Findings and interpretations of existing studies and relevant theory.
Data Collection: Purposive sampling	**Data Collection:** Purposive sampling.	**Data Collection:** Convenience sampling.
Process: Constructing interpretations	**Process:** Constructing interpretations.	**Process:** Bridging summaries.
Product: Conceptual maps and interpretations	**Product:** Interpretations across case studies.	**Product:** Logical rationalizations.
Steps Description [5]	**Steps Description [19]**	**Steps Description [25]**
Extract data: Extract data from the primary studies, including bibliographical information, aims, context, and results.	**Data Reduction:** Process of selecting, focusing, simplifying, abstracting and transforming the results from studies.	**Developing a theoretical model of how the interventions work, why and for whom**: Inform decisions about the review question and what types of studies to review.
Code data: Identify and code interesting concepts, categories, findings, and results in a systematic fashion across the entire data set.	**Data Display:** A display is an organized, compressed assembly of information that permits conclusion drawing and action using a "tool-box". The "tool-box" includes un-ordered, site-ordered, and time-ordered meta-matrices, scatterplots, and cause and effects graphs or networks	**Developing a preliminary synthesis:** To organize findings from included studies to: describe patterns across the studies in terms of the direction or size of effects; to identify and list the facilitators and barriers to implementation reported.
Translate codes into themes, sub-themes, and higher order themes.		
Create a model of higher-order themes: Explore relationships between themes and create a model of higher-order themes.	**Conclusion Drawing and Verification:** From the start of data collection, the qualitative analyst is beginning to decide what things mean – is noting regularities, patterns, explanations, possible configurations, causal flows and propositions. Conclusions are also verified as the researcher proceeds, The meanings emerging from the data have to be tested for their plausibility, their sturdiness, their "confirmability" – that is, their validity.	**Exploring relationships in the data**: To consider the factors that might explain any differences in direction and size of effect or facilitators and/or barriers to successful implementation across the included studies; To understand how and why interventions have an effect.
Assess the trustworthiness of the synthesis: Assess the trustworthiness of the interpretations leading up to the thematic synthesis.		
		Assessing the robustness of the synthesis product: To provide an assessment of the strength of the evidence for drawing and generalizing conclusions to different population groups and/or contexts.

Data reduction is the identification of items of evidence in the primary studies. (It is worth noting that the major data reduction is conducted in the analyses in the primary studies themselves). Data is then clustered into meta-matrices and time-ordered displays, which are used to draw conclusions from the synthesized studies. The use of matrices and tables facilitates the comparison of the cases and areas of

agreement or disagreement across cases. Miles and Huberman classify cross-case clustering approaches in variable-oriented or case-oriented. In variable-oriented approaches, variables identified in the cases take center stage, that is, the inner-dynamic of the case is replaced with a search for patterns and themes that cut across the cases; the pressure is put on the researcher in terms of interpreting the answers so that they can be reduced to variables. In case-oriented approaches, commonalities across multiple instances of a phenomenon may contribute to conditional generalizations thought formation of types or families of studies. One advantage of the method is the transparency that the data-matrices allow to the process of synthesis. One disadvantage is that it may lead to conclusions of the abstracts levels of the variables and cases without considering the whole context of the studies.

Narrative synthesis refers to an approach of synthesis that relies primarily on the use of words and text to condense and explain the findings of the synthesis. Whilst narrative synthesis can involve the manipulation of statistical data, the defining characteristic is that it adopts a textual approach to the process of synthesis to 'tell the story' of the findings from the included studies [25][26]. As used here 'narrative synthesis' refers to a process of synthesis focusing on a wide range of questions, not only those relating to the effectiveness of a particular intervention. It is a general approach within which a wide range of specific methods of synthesis can be used. Popay et al. [25] define four main elements of a narrative synthesis process (Table 2): *theory development, development of a preliminary synthesis, exploring relationships in the data*, and *testing the robustness of the synthesis*. Around 20% of the synthesis methods in systematic reviews in SE can be classified as narrative synthesis [4]. However, none of these systematic reviews are explicit about which approach was followed. The lack of transparency and lack of an authoritative body of knowledge as well as the lack of reliable and rigorous techniques are among the drawbacks of the approach. The data collection is also a point of debate as there is not a systematic defined criterion to choose the data and it is usually based on the convenience of the analyst. The framework by Popay et al. [25] has the potential to produce more transparent and more sophisticated narrative syntheses if they start to be adopted in SE.

III. WORKED EXAMPLE

To investigate the research questions posed in this paper, we conducted two independent syntheses of two published case studies (on trust in outsourcing relationships) [2][24]. We defined a common synthesis goal and ran one synthesis in Sweden (using cross-case analysis) and the other in Norway (using thematic synthesis). These two syntheses were then compared to a third, already conducted narrative synthesis of the two case studies. The common goal of the syntheses was to:

Understand factors of trust in outsourcing relationships.

This is a knowledge support goal and not a decision support goal [1][26]. A synthesis directed at knowledge support will typically bring together and synthesize research evidence on a particular topic aiming at creating new knowledge on the topic. We identified two papers that could help us to fulfill our goal: Oza et al. (Oza et al. study) [24] and Babar et al. (Babar et al. study) [2]. They were selected based on their relatively high homogeneity, investigating very similar research questions, from a similar perspective, although in two different contexts, two years apart, and with two different sets of researchers. Preliminary versions of both studies were published at the EASE conference in 2005 and 2006 [22][23], respectively. At the 2006 conference, the similarity between the two studies were observed, leading to the latter study being extended with a narrative synthesis between the two, when expanded into a journal version [2]. Interestingly enough, only one of the papers was included in an SLR of global software engineering, despite their similarity [30].

The Oza et al. study, was based on interviews of 18 software development practitioners in India [24], while the Babar et al. study was based on interviews of 12 Vietnamese practitioners developing software for Far Eastern, European, and American clients [2].

The goal of the Oza et al. study was to investigate the following research questions:
    i) What are the critical factors to achieving trust initially in an outsourcing relationship?
    ii) What are the critical factors to maintaining trust in an established outsourcing relationship?
The goal of the Babar et al. study was to investigate what factors are important for:
    i) Establishing trust in off-shore software outsourcing relationships, and;
    ii) Maintaining and strengthening trust in offshore software outsourcing relationships.

A secondary goal of the journal version of the study by Babar et al. was to compare their results with Oza et al. (the first study). This comparison was performed through narrative synthesis. We decided to not read the narrative synthesis before we had performed our own syntheses. For the data collection, the Oza et al. study used standardized open-ended interviews to collect qualitative data. Babar et al. used semi-structured interviews based on a modified version of the questionnaire developed and used by Oza et al. Both studies used qualitative data analysis approaches for reaching conclusions. Both studies also have their own definitions for each factor of trust. These definitions are reproduced in Tables 3 and 4.

In the following, we describe how we performed the syntheses and what were the results from each synthesis process: thematic, cross-case, and narrative synthesis.

TABLE 3 - DEFINITION OF TRUST AS DEFINED BY OZA ET AL. [24]

Initial and Maintaining Trust Factors

Trust	Trust is investigated at two levels: (1) initial trust when outsourcing relationship has not started and (2) after the relationship has started
Role (importance)	Refers to the role of trust in outsourcing relationships (in vendor's opinion). It also looks at the important factors to achieve trust from the client
Initial trust	How vendor achieves first time (initial) trust when outsourcing engagement is in the prospective stage or has just started
References	Vendor's opinion about how references from their previous clients is useful to them in achieving trust from the prospective client
Experience	How vendor's experience in the outsourcing industry helps to gain trust from the client
Reputation	Vendor's opinion about how certifications from international companies, successful project histories and other previous achievements lead to a good reputation of the company and in turn if becomes useful in achieving trust from the prospective client
Client visits	Vendor's views about the client visits to their premises, how it can help gaining trust from the client
People background	Skilled workforce available to the vendor and their backgrounds and credentials which help to the success of outsourcing
Investment	Vendors views on his willingness to invest in the outsourcing project through the company's financial strength, allocations, etc. to make the project successful
Trust (ongoing)	Investigation of trust factors when outsourcing relationship has already started (ongoing)
Transparency	How vendor's transparent actions/outcomes can help to gain more trust. It also refers how client is transparent in sharing the necessary information in outsourcing engagement
Demonstrability	Demonstrability of the work done and articulating the facts in a right manner, which can help in gaining trust
Honesty	How vendor's honesty assist in gaining trust, honesty here is referred in terms of presenting the real facts about the outsourced work, reacting proactively if something is wrong, and performing honestly with the client in terms of outsourcing operations
Process	Processes followed by the vendor to complete the outsourced work successfully. Some vendors also emphasized process driven approach to gain trust
Commitment	How commitment to the outsourced work can help vendor to gain trust from the client. It also comprises that in vendors opinion, it is better to under commit and than over deliver rather than doing over commitment and under deliver which can be destructing in gaining the trust
Communication	How communication can help maintaining trust with the clients
Cooperation	For outsourcing success, how it is useful to cooperate by contributing the necessary inputs (from the client and the vendor side). How both companies can support each other in tough situations
Consistency	How consistently you can maintain trust from the client. How consistently vendor can deliver the outsourced services/work successfully, how consistently vendor can maintain trust from the client
Understanding	Understanding between clients and vendors in transacting with each other
Confidentiality	Many outsourced services/products also carries sensitive information which should be treated with strict confidentiality by the vendor and they should be able to demonstrate that
Performance	You have to perform the work to gain the trust, it is based on performance

TABLE 4 - FACTORS IMPORTANT TO ESTABLISH AND MAINTAIN TRUST RELATIONSHIP, AS DEFINED BY BABAR ET AL. [2]

Initial Trust Factors	
Cultural understanding	How knowledge of the norms, beliefs, business ethos, and skill in the native language of potential clients helps vendors achieve trusts
Creditability	How references, certifications, previous experiences help to gain trust from clients
Capabilities	How technology, people and management capabilities of vendors help to gain trust from clients
Pilot project Performance	How performance of pilot projects help to gain trust from clients
Personal visits	How visits by clients to vendors' development facilities help to gain trust from clients
Investment	How investments of vendors in people, technologies and infrastructure help to gain trust from clients
Maintaining Trust Factors	
Communication	How effectiveness of communication with clients (maybe in clients' native language) help to maintain the trusts
Cultural understanding	How knowledge of the norms, beliefs, business ethos, and skill in the native language of potential clients helps vendors achieve trusts
Capabilities	How technology, people and management capabilities of vendors help to gain trust from clients
Contract conformance	How observation of all clauses in business agreement, protection of intellectual properties help to gain trusts from clients
Quality	How quality of delivered products help to maintain trusts with clients
Timely delivery	How adherence to development schedule helps to maintain trusts with clients
Development processes	How processes followed in the outsourced development help to gain trusts from clients
Managing expectations	How to raise fulfillable expectation to clients help to maintain trusts
Personal relationships	How personal relationships between clients and vendors at different levels of management and development team help to maintain trusts with clients
Performance	How performance (productivity/effectiveness) of staff in carrying out the projects help to maintain trusts with clients

A. Thematic Synthesis

The thematic synthesis followed the steps and checklist proposed by Cruzes and Dybå [5] (see also Table 2), and was performed by the Norwegian team. Five steps were performed (as described in Figure 1): initial reading of data/text (extraction), identification of specific segments of text, labeling of segments of text (coding), translation of codes into themes, creation of the model and assessment of the trustworthiness of the model.
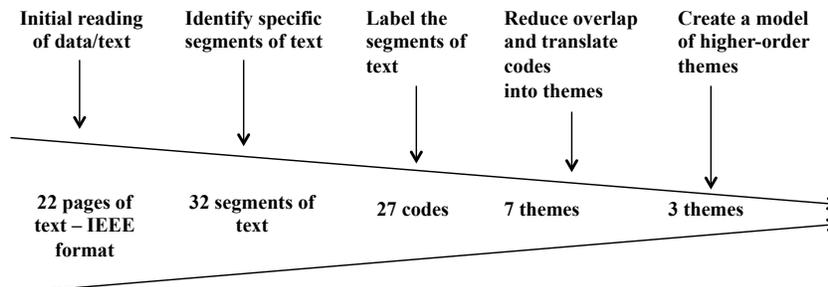

FIGURE 1 - PROCESS OF THEMATIC SYNTHESIS FOLLOWED IN THE WORKED EXAMPLE (ADAPTED FROM [5])

THE EXTRACTION OF THE DATA CONSISTED OF THE PUBLICATIONS' DETAILS (AUTHORS, TITLE AND PUBLICATION YEAR), THE CONTEXT (GEOGRAPHY), AND THE STUDY RESULTS (FACTORS OF TRUST IN OUTSOURCING RELATIONSHIPS). WE USED NVIVO TO HELP ON THE IDENTIFICATION OF THE SEGMENTS OF TEXT CONTAINING REFERENCES TO FACTORS OF TRUST IN THE TWO PAPERS (TABLE 3 AND

Table 4). The coding was also done using NVivo and consulting the list of definitions of each factor as used by the authors of each paper. As shown in Figure 1, we extracted 32 segments of text from the 22 pages of the two papers (references in NVivo). From these segments, 27 codes were abstracted considering the commonalities and differences on the definitions and the text where the definitions were quoted (as shown in Figure 2). For each code, it is possible to retrieve the definition given by each paper

to that concept and also get a link to original text where the code came from, in Figure 3 the communication node is shown, it has two segments of text that specifically describe communication as a factor of trust in outsourcing relationships: Oza et al. defined it as: "How communication can help maintaining trust with the clients," while Babar et al. defined it as: "How effectiveness of communication with clients (maybe in clients' native language) help to maintain the trust". As we can see, the definitions of communication in the two papers differ slightly, and in these cases we needed to create a new definition that would encompass both definitions.

We reduced overlap and translated the 27 codes into the following seven themes: Commitment, Communication, Development Process, Investments in People, Technologies and Infrastructure, Reputation, Team Member Skills, and Team Performance (as shown in Figure 2). Now, Communication (Figure 4) is a theme composed of four codes: transparency, personal relationships, honesty, and communication. The definitions and the quotes from these codes were all related to the more abstract concept (or theme) 'communication', which we defined as: "How a regular process by which information is exchanged between individuals through a common system of symbols, signs, or behavior can help maintaining trust with the clients."

Finally, we created a model of higher-order themes where we mapped the seven themes into three higher order themes: Initial Trust, Maintain Trust, and Initiating and Maintaining Trust. On these themes the seven previous mentioned themes were organized. The final concept map is the one shown in Figure 5. For each entity of the mind map there is some information associated to it: Definitions from the paper, references to text backing up these definitions, a note showing in which paper the factor appeared, and for each of the seven main themes there is also a conclusion and a definition associated with it, as shown in Table 5. The strength of the conclusion is based on the number of times mentioned by the interviewees in each study.

The trustworthiness of the model was a straightforward activity because we had only two papers to relate to, therefore all the codes and references could be easily mapped back to the original papers. Besides, we were two researchers doing the work and assessing every step of the process.

FIGURE 2 - CODES IN NVIVO FOR THE THEMATIC SYNTHESIS

FIGURE 3 - COMMUNICATION AS A NODE IN THE SYNTHESIS – ON THE RIGHT SIDE OF THE FIGURE THERE IS THE MINDMAP OF THE THEMATIC SYNTHESIS.AND ON THE LEFT THE NOTES (DEFINITIONS AS DESCRIBED IN THE PAPERS AND THE REFERENCE ON THE TEXT DESCRIBING THE EVIDENCE FOUND ON COMMUNICATION ON BOTH PAPER
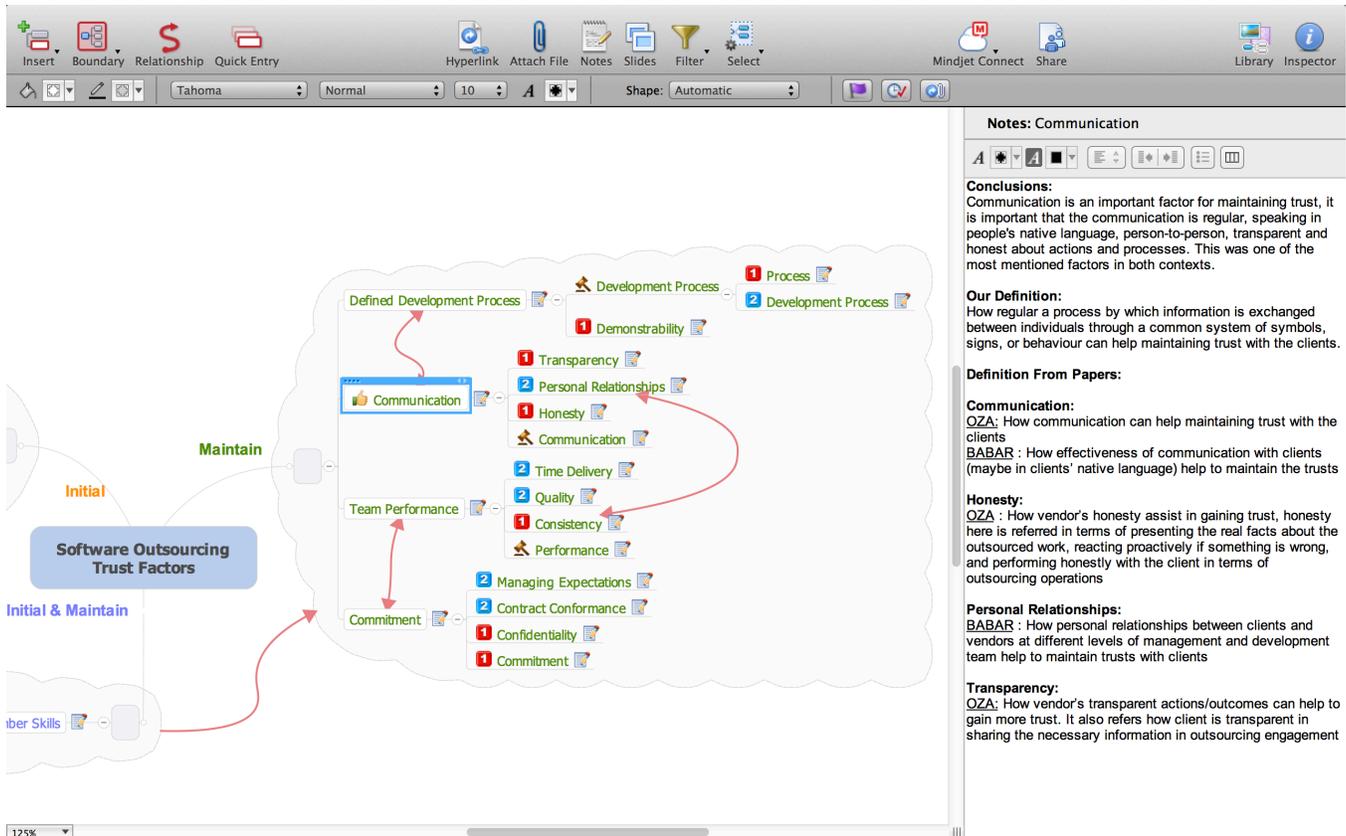
FIGURE 4 - COMMUNICATION AS A THEME IN THE THEMATIC SYNTHESIS. THE RIGHT SIDE OF THE FIGURE SHOWS THE MINDMAP OF THE THEMATIC SYNTHESIS AND THE LEFT SIDES SHOWS THE NOTES FROM THE RESEARCHERS ON THE THEME COMMUNICATION THAT CONSISTS OF FOUR NODES (TRANSPARENCY, PERSONAL RELATIONSHIPS, HONESTY AND COMMUNICATION)
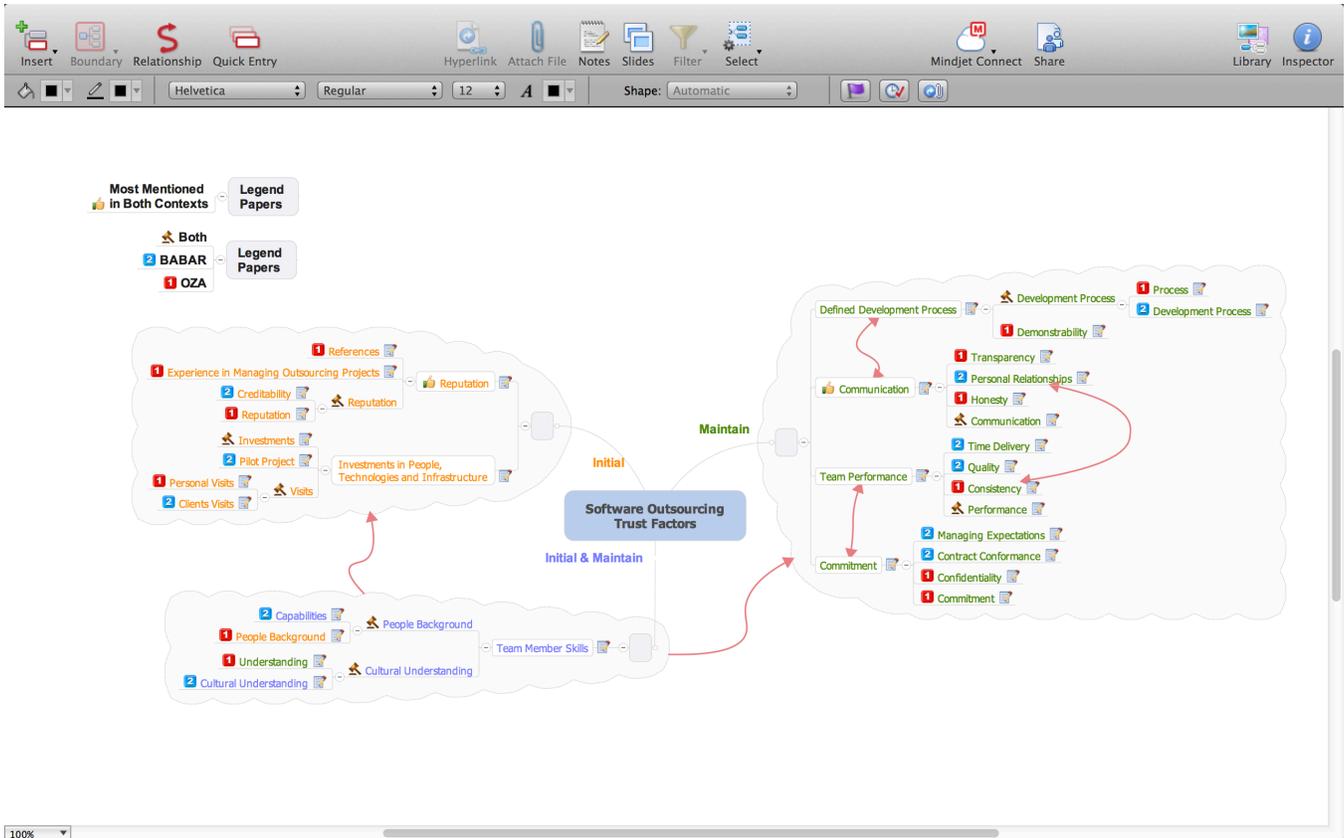
FIGURE 5 - CONCEPT MAP FROM THE THEMATIC SYNTHESIS. THE MAP DESCRIBES THE FINAL THEMATIC MAP OF THE SYNTHESIS; SOME FACTORS WERE FOUND ONLY IN ONE PAPER (1 FOR OZA AND 2 FOR BABAR) OR BOTH. THE MOST MENTIONED OF ALL BY THE INTERVIEWS ARE ALSO MARKED. A RED ARROW SHOWS THAT THERE IS A RELATIONSHIP BETWEEN THE NODES OR THEMES.

TABLE 5 - THEMES DEFINITIONS AND CONCLUSIONS

| Higher-Order Theme	Theme	Definition	Conclusion
Main factors for initiating trust with vendors	Reputation	How references, certifications and previous experiences help to gain trust from clients	Reputation is an important factor for initial trust in a relationship of outsourcing. Ways of showing reputation include references, previous experience and certifications. This factor was strongly mentioned in both contexts.
	Investments in people, technologies and infrastructure.	How the perceived investments of the vendor in people, technologies and infrastructure help to gain trust from clients	It is important for the clients at a first moment that the vendors shows that they are investing in the relationship, this can be done by visits to the client, piloting or allocating personnel for the project. In both contexts this factor was mentioned but not strongly mentioned.
Main factors for maintaining trust with vendors	Defined development process	How processes explicitly followed in the outsourced development help to gain trust from clients.	Defined processes that are shared with the clients are important factors for maintaining trust. They are also important to maintain a good relationship and communication with the clients. This factor had about the same strength in both contexts.
	Communication	How regular a process by which information is exchanged between individuals through a common system of symbols, signs, or behavior can help maintaining trust with the clients.	Communication is an important factor for maintaining trust, it is important that the communication is regular, speaking in people's native language, person-to-person, transparent and honest about actions and processes. This was one of the most mentioned factors in both contexts.
	Team Performance	How consistent timely and quality delivery of vendors help to maintain trust with clients.	Teams have to deliver in time and with good quality consistently to maintain the trust with the clients. This is mentioned by about half of the interviewees in both contexts.
	Commitment	How confidentiality, contract conformance and management of expectations to the outsourced work can help vendor to maintain trust.	Commitment to the outsourced work is an equally important factor for managing trust in an outsourcing relationship in both contexts.
Factor of Trust for both initiating and maintaining trust with vendors	Team Member Skills	Vendors' skilled workforce with background and credentials that help to the success of outsourcing.	The skills of the workforce are an important factor for trust in outsourcing relationships. Mentioned skills are related to technical capability, people capability and management capability, but these skills are mostly mentioned as initial factors for trust, although also mentioned as a maintain factor in the Vietnamese context. Cultural Understanding was mentioned mostly in the Vietnamese context, appearing as strong factor for both the initial trust and maintenance of trust.

B. Cross-Case Analysis

The cross-case analysis method is not a prescriptive step-by-step procedure; instead it offers a high-level three-step method, and a "tool-box" of cross-case displays, primarily matrices, to organize that data by variable and/or by case. The process is most clearly presented in [19], while the toolbox is introduced in [20]. Also remember that the method is originally presented as a synthesis method *within* a multi-case study, while we here use it to synthesize *across* two single-case studies. The method has three major steps: 1) data reduction, 2) data display, and 3) conclusion drawing and verification (See Table 2). As in any qualitative analysis method, the steps are iterated during the analysis to reach the final conclusion.

In our worked example, the major part of *data reduction step* was conducted already in the analysis in the primary studies. The synthesis focused on reduction of the material into data in both primary studies, which were stated to have an impact on trust in the outsourcing relationships. Since we only synthesized evidence from two journal papers, which were quite condensed and homogenous, we could get an overview of the papers in their raw formats, and tagged data directly in printouts of the papers. If more studies are synthesized, or if less homogenous studies are synthesized, data have to be stored in, for example, NVivo to allow easier navigation in the data.

TABLE 6 – REPORTED CHARACTERISTICS IN THE PAPERS

Aspect	Oza et al. study	Babar et al. study
Goal	RQ 1: What are the critical factors to achieving trust initially in an outsourcing relationship? RQ2: What are the critical factors to maintaining trust in an established outsourcing relationship?	Identify factors for • establishing trust in off-shore outsourcing relationships • maintaining and strengthening trust in off-shore outsourcing relationships
Target	Indian software developers	Vietnamese software developers
Target culture	India – influenced by Britain	Vietnam – influenced by France
Cases	18 companies, 18 interviewees	8 companies, 12 interviewees
Maturity	CMMI 4 or 5, 14 out of 18 > 10 years experience Size: 1000-5000-10000-15000	Mostly young with 5-8 years of experience, companies 5-10 years
Methodological framework	Yin [32]	Yin [32]
Data collection	Standardized open-ended interviews	Semi-structured interviews
Data analysis	Grounded theory: Open coding Inter-rater reliability tests	Content analysis Frequency analysis
Key concepts	Trust: - willingness to be vulnerable - (positive) expectations	Trust: positive expectations [23]. - calculus-based - knowledge-based - identification-based - performance-based

The data we derived from the papers were of two kinds: 1) characteristics of the case studies, and 2) factors of trust in outsourcing relationships. Most of these data were presented under easily found section headings in both primary studies, and hence straightforward to derive. All data about the characteristics available in the papers were collected, including facts about the studies' goals, the companies and interviewees participating in the study, the research methods used for data collection and analysis, and the theoretical frame of reference for the "trust" concept. These data are tabulated in Table 6. The aspects of the case characteristics originate from the aspects presented in the primary studies. It can be worth noting that the concepts of maturity and size of the companies are measured using different characteristics in the two cases; by CMMI level, employer experience, and size in the Oza et al. study, and by employer experience and company age in the Babar el al. study.

THE TRUST FACTORS IDENTIFIED IN THE TWO STUDIES WERE GIVEN IDENTIFICATION TAGS FOR INITIATION AND MAINTAINING FACTORS, RESPECTIVELY (E.G. MB4 IS THE 4TH MAINTAINING FACTOR LISTED BY BABAR ET AL.). SOME FACTORS INCLUDED SUB-CHARACTERISTICS, WHICH ARE RATHER SPECIFIC (E.G. IB2 CREDITABILITY HAD ONE SUB-COMPONENT, IB2.1 REFERENCES). AS AN ACT OF *DATA DISPLAY*, THESE WERE PRESENTED IN AN UN-ORDERED MATRIX (WHICH WAS LATER ORDERED, SEE TABLE 7), ONE COLUMN FOR EACH STUDY, WITHOUT ANY CONSIDERATION OF THE SEMANTICS OF THE TERMS. IN THE CURRENT SYNTHESIS, THE PROCESS WAS PRETTY STRAIGHTFORWARD, AS THE OZA ET AL. STUDY HAD THE FACTORS COLLECTED IN AN APPENDIX, PRESENTING THEIR CODEBOOK, AND THE BABAR ET AL. STUDY PRESENTED THEM IN TWO TABLES (WHICH WE REPRODUCED IN TABLE 3 AND

Table 4). One additional factor was mentioned in their text, but not handled as a factor in the tables, but we thought it was important to add it in our matrix (Factor IO7 Representativeness).

The next step in the analysis of trust factors was a new act of *data reduction* to analyze the semantics of the identified factors. (Remember that the three analysis steps of cross-case analysis are not sequential, but iterative.) We identified the synonyms and hyponyms based on the definitions of the factors in each case, and their presentation in context. We rearranged the factors table, based on the semantic meaning of terms in the two studies. Three of the identified factors had the same meaning in both studies (IO4/IB5 personal visits, IO6/IB6 investments, and MO4/MB10 performance). Two factors had different terms for approximately the same underlying concepts (IO5 people background = IB3 capabilities; MO6 commitment = MB8 managing expectations). In five cases, one factor in one of the studies had two,

three, or four hyponym factors in the other study (for example, IO1 references, IO2 experience, and IO3 reputation in the Oza et al. study, corresponds together to the term IB2 creditability in the Babar et al. study, and are hence hyponyms of creditability). Eight factors appear in one study only and have no correspondent in the other (IO7, IB1, IB4, MB2, MB3, MB4, MB9, and MO9).

TABLE 7 - RELATED FACTORS IN THE TWO STUDIES, TAGGED XYN, WHERE X={I,M} FOR INITIATION AND MAINTAINING FACTORS, Y={O,B} FOR OZA AND BABAR STUDIES, AND N IS AN ORDER NUMBER FOR THE FACTOR FOUND IN THE PRIMARIY STUDIES.

	Oza et al study	Frequency	Babar et al study	Frequency
Initial factors				
	IO1 References	14	IB2 Creditability IB2.1 References IB2.2 CMM level IB2.3 Agent	11
	IO2 Experience	9		
	IO3 Reputation IO3.1 CMM level	6		
	IO5 People background	4	IB3 Capabilities	9
	IO4 Personal visits	5	IB5 Personal visits IB5.1 Technical staff IB5.2 Managerial staff IB5.3 Move staff	7
			IB1 Cultural understanding	12
	IO6 Investments	4	IB6 Investments	6
			IB4 Pilot project performance	8
	IO7 Representatives	>1		
Maintaining factors				
	MO1 Transparency MO1.1 Project tool MO1.2 Real actions	10	MB1 Communication	12
	MO3 Honesty	9		
	MO5 Communication	8		
	MO8 Understanding	7		
	MO2 Demonstrability	9	MB5 Quality	10
			MB6 Timely delivery	9
	MO4 Process	9	MB7 Development processes	9
	MO7 Consistency	7		
	MO6 Commitment	8	MB8 Managing expectations	8
			MB2 Cultural understanding	12
			MB3 Capabilities	11
			MB4 Contract conformance	10
	MO10 Performance	4	MB10 Performance	6
			MB9 Personal relationships	7
	MO9 Confidentiality	4		

Both studies used quasi-statistics in their analysis, which is a doubtful practice, if interpreted wrongly [29]. However, it may be used to bring forward the most frequently identified factors, and hence we site-ordered meta-matrices based on the sum of the term frequency in the two studies, resulting in Table 7. The ordered table shows that IB2 creditability (Babar), and its hyponyms IO1 references, IO2 experience and IO3 reputation (Oza) are the most frequently mentioned factor for trust establishment, while MB1 communication (Babar) and its hyponyms MO5 communication, MO3 honesty, MO1 transparency and MO8 understanding (Oza) are the most frequently mentioned maintaining factors.

It is also clear from the matrix that IB1 cultural understanding and IB4 pilot project performance are identified as trust establishing factors only in the Babar et al. study, and only the Oza et al. study mentions the importance of the representatives sent forward to represent the company (IO7). Further, MB2 cultural understanding, MB3 capabilities, and MB4 contract conformance, are identified as trust maintaining factors only in the Babar et al. study, while MO9 confidentiality is only identified as a maintaining factor in the Oza et al. study.

The synthesis activity of *conclusion drawing* was to identify the relations stated between the factors in each of the primary studies and express them in a graph, see Figure 6. These relations were expressed qualitatively in the analysis text. For example, Oza et al. stated that: *"vendors also consider their market reputation a critical factor to gain trust initially. This reputation was also part of the good references and long experience of software outsourcing. Reputation building was also reported to be based on the CMM and other quality certifications…"* This statement was captured in the top-left corner of the graph in Figure 6, where factors IO1 References, IO2 Experience, and the IO3.1 CMM level contribute to factor IO3 reputation, which is a subset of the factor IB2 creditability.

For the maintenance of trust, Oza et al continue: *"the majority of vendors identified* transparency *as a critical factor to maintain trust with the clients. Vendors identified transparency in undertaking the process, communicating with the client and showing outcomes of the project. One vendor commented: We have a project office tool. With the use of this tool clients can view each employee's timesheet information on a daily basis and the work status. We are always happy to make it open to the customer, if he wants, he can get it. When you open the whole system process to somebody it gives lots of confidence and they trust you."*. This was interpreted as having a project office tool (MO1.1), contributes to MO1 transparency, which in turn contributes to maintaining trust in the outsourcing relationship.

These relations were listed for each study and then displayed in graphs, see an example in Figure 6. Note that only factors with more than one relation are drawn in the graph for visibility reasons, i.e. factors stated to have only a single relation to trust are not included in the graph. Relations originating from the Oza et al. study have dashed lines, and relations from the Babar et al. study have solid lines.

The *conclusion drawing and verification step* involved refinement of the above steps. We phrased condensed summaries of each of the papers' views, for example, on the trust factors for maintaining trust as follows:

- Oza et al. present different aspects of *transparency* as critical success factors. Trust grows when you demonstrate that you have nothing to hide. This is well in line with their definition of trust as "willingness to be vulnerable". Examples of transparency are 1) having a project office tool, where the client can monitor all project data, 2) backing up statements and promises with real actions, 3) being honest and not hiding anything to the client, 4) using processes as a framework to relate the progress to.
- Babar et al. focus on *communication* and *cultural understanding* in their analysis. Communication is the basis for building and maintaining trust, both formal and informal relationships. The communication leads to cultural communication through mutual exchange visits, which in turn improves trust. Hence, the communication is assumed to have one direct and one indirect impact on trust.

The cross-case analysis does not reveal any contradictions between the two studies with respect to trust factors, meaning that one did not state the opposite of what the other states. The frequency ranking is also very much the same in the two studies. However, they put different emphasis on the factors when they qualitatively discuss a few key ones.
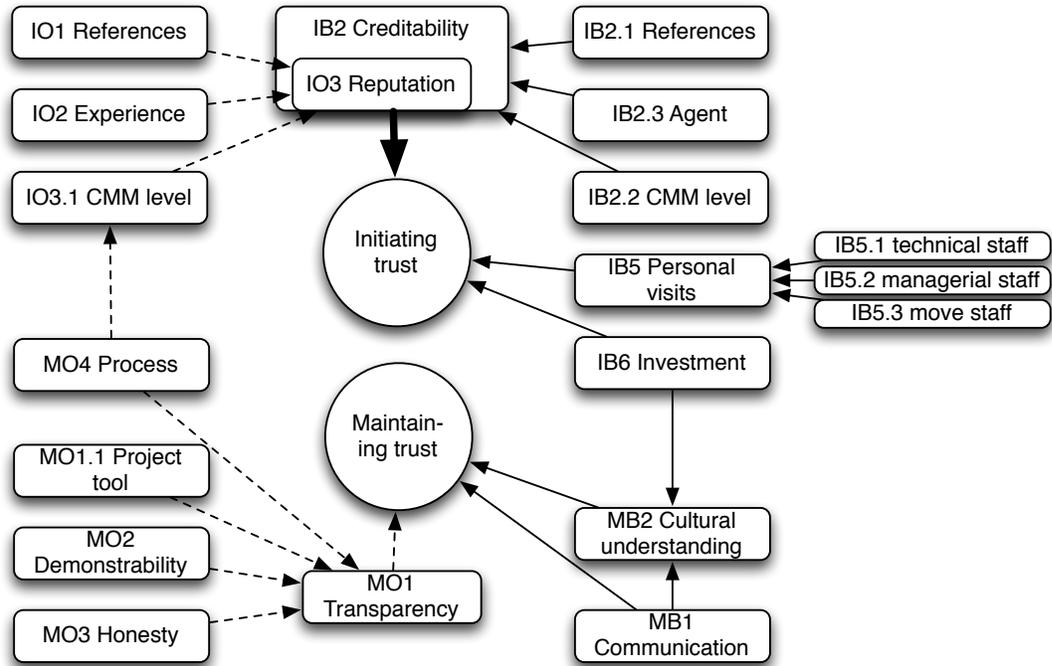
FIGURE 6 - GRAPH FROM CROSS-CASE ANALYSIS. THE DASHED LINES ARE RELATIONS FROM OZA ET AL WHILE SOLID LINES COME FROM BABAR ET AL.

c. Narrative Synthesis

Apart from understanding the perspectives of Vietnamese vendors (Babar et al. study) with regard to trust between clients and vendors in the context of off-shore software development outsourcing, another objective of Babar et al.'s research was to compare the Vietnamese practitioners' views with their Indian counterparts [2]. In this section, we present their comparative analysis of factors identified as important in establishing and maintaining trust relationships by the Vietnamese and Indian practitioners, which was made by the authors of the second paper, Babar et al. The data for Indian practitioners' views were taken from the study reported by Oza et al. (2005) and the data and narrative synthesis for the Vietnamese vendors is fully described in the paper by Babar et al (2007). Babar et al. structured their synthesis in two main sections: Establishing trust and maintaining trust. The authors structured the narrative based on comparing the results from the two settings.

For establishing trust Babar et al synthesize the findings as:

"For factors perceived as important for establishing initial trust between clients and vendors, Indian and Vietnamese practitioners seem to agree on only two factors: 'Client visits' and 'Investment'. A semantic analysis of the factors identified by both groups revealed that Indian practitioners mentioned 'Customer references', 'Experience in outsourcing', and 'Reputation' as important factors, which can be semantically considered quite close to 'creditability', a factor mentioned by the Vietnamese practitioners. Vietnamese practitioners also agree on the importance of some other factors (such as references, experience, reputation, and creditability) in gaining a client's trust initially."

Babar et al. also give some of their perceptions on similarities and perceptions in the results for the main factors perceived as similar for the two settings, as for instance in the following part of the narrative synthesis:

"Vietnamese companies are relatively new in the software outsourcing business and have

relatively few significant customers, whose references they can use to gain the trust of new customers. However, Indian companies, being veterans of the software outsourcing business, do not face this situation, as they are able to use the references of large multi-national companies, who have outsourced their software development to Indian companies. One of the interviewees of study conducted by Oza et al. emphasized the role of customer references in gaining initial trust in these words: 'References help us to a great extent, if I pull out a long customer list then people will say see they are working with scores of such companies, they can work for us.'"

An interesting part of the narrative synthesis was that they also discussed the main differences in the findings for initial trust:

"One significant difference between Vietnamese (Babar et al. study) and Indian practitioners' views (Oza et al. study) is the role of 'cultural understanding' in establishing a trust relationship. As we discussed previously that Vietnamese vendors consider the understanding of a client's culture as a critical factor in gaining initial trust. However, it appears that their Indian counterparts do not perceive the cultural understanding being of any importance, as they do not mention anything about the cultural understanding at all.

The narrative synthesis approach was particularly useful in describing the differences and in making explicit the diversity in the study context. For example, Babar et al., when explaining the differences between the two contexts in terms of the factor "cultural understanding" wrote:

"One explanation for Indian practitioners not mentioning cultural understanding can be their familiarity with the culture and language of their major clients, usually Americans. Understanding the written and spoken language (English), coupled with a strong linkage to Western countries through expatriates has been widely cited for Indian companies' success in attracting outsourced contracts (Carmel, 2003a). On the other hand, Vietnam has its own language and business ethos, which necessitates learning the languages and gaining a cultural understanding of their clients. That is why Vietnamese practitioners view 'cultural understanding' as one of the most important factors in gaining a client's trust for establishing long-term relationships. This is particularly true for their Japanese clients, who because of the strong uncertainty avoidance nature of their culture (Hofstede, 1980) prefer long term business relationships to gaining short term benefits."

The same structure of the narrative synthesis, similarities and differences, was used for establishing trust and for maintaining trust. Babar et al. summarize the main similarities in the findings of establishing trust as:

"Both groups agree that processes, communication and performance are important factors for maintaining trust between clients and vendors. We also observe that both Indian and Vietnamese practitioners also seem to agree on some other factors as important in maintaining trust, though the description of these factors is lexically different. For example, Indian practitioners mentioned 'Honesty', 'Commitment', 'Confidentiality', 'Cooperation', and 'Understanding', while the Vietnamese used constructs like 'Contract conformance', 'Managing expectations', and 'Personal relationship' to describe similar factors. For instance, Oza et al. quoted one of the Indian practitioners emphasizing the importance of 'Honesty' in maintaining trust in ongoing relationships in following words:

> *'You have to be upfront and honest with your client. You should not hide anything from him, whether it is good or bad, whether it is going to earn you a flack for that moment. This is very important for the long lasting relationship and to achieve*

trust.'

Compared with the above-mentioned views of the Indian practitioner on the role of Honesty, Vietnamese practitioners described the requirement of being Honest and open to clients to gain and maintain their trust in terms of contract conformance and managing expectations. One Vietnamese interviewee described the importance of being honest and upfront in the following words:

'It is very important to demonstrate to your clients that there are certain measures in place to make all project team members aware of the criticality of the conformance with the contractual obligations; it might be conformance with the non-disclosure agreement (NDA) or keeping commitment to the deadlines and budgetary limits. Be upfront in explaining what can or cannot be done within the given time and budget. Change requests must be monitored carefully and customers should be taken into confidence if something goes against the plan, for example, key members of a project team leaving the company.'"

The main difference perceived in the findings of factors in the maintenance of trust in outsource relationships is again the importance of the cultural understanding. Babar et al. conclude that:

"The views of Vietnamese and Indian practitioners again differ on the importance of 'Cultural understanding'. The Vietnamese considered this factor also vital to maintaining trust in ongoing relationships, while their Indian counterparts did not mention this factor again. Thus, a major difference between the Indian and the Vietnamese practitioners' views is that the former seem to consider factors related to business process more important, while the latter not only realize the importance of business process related factors but also recognize the vital role of cultural understanding and personal relationships in maintaining a trust relationship. These two factors are considered very important for successful business partnership by Asian clients in general and by the Japanese in particular, who are major software outsourcing clients of Vietnamese vendors."

To summarize, Babar et al, concludes the synthesis saying:

"Another significant point revealed by this comparative analysis is that the Indian practitioners identified entirely different factors that they considered important for establishing trust and maintaining trust, while there are a few factors that the Vietnamese practitioners considered important for both the establishment and maintenance of trust in software outsourcing relationship. For example, the Vietnamese practitioners described cultural understanding and capabilities as important factors for both establishing and maintaining trust."

IV. DISCUSSION

In this section we discuss and provide answers to the research questions of this study. Our aim is to demonstrate the similarities and differences of results and conclusions when applying different methods of synthesis and to discuss the challenges of synthesizing evidence from reported case studies in SE. Our main research questions were:

i) What are the differences in the results when using narrative, cross-case or thematic synthesis of case studies evidence?

ii) What are the main challenges of performing case studies synthesis?

a) Comparison of results from Methods of Synthesis

For the purpose of this paper three of the most relevant methods are compared: thematic synthesis, cross-case analysis, and narrative synthesis. This comparison is performed based on a worked example

as shown in the previous section. Before doing an analysis of the comparison of the syntheses, we would like to note the thematic synthesis and the cross-case analysis were performed by researchers that were not involved in any of the two primary studies. However, the narrative synthesis was performed by the three authors of the second study (Babar et al. study), therefore, they also had access to the raw data of the second study, which may have given them the opportunity to go deeper in their synthesis. Whether the findings might be different with independent researchers or not, is per se a researchable question. It is clear from the example that the facts are taken into the narrative analysis, which were not in the original study (cultural origin, maturity of insourcing in the country), which the more structured methods tried to avoid. However, while synthesizing the results from the two papers, both teams, Norwegian (thematic synthesis) and Swedish (cross-case analysis), found that the quotes inserted in the papers were not enough to be totally confident that we were synthesizing the papers at the right level of abstraction and granularity.

In our example, the primary studies had the same goals and methodological framework. The main variations were the target culture (India vs. Vietnam) and the research groups. There was a temporal variation in the sense that Babar et al.'s study was run based on Oza et al.'s previous paper and results. There is hence a threat that Babar et al.'s results may be influenced by Oza et al.'s results. But they also added two important variations: definitions and target cultures. The terminology and definitions are partly different; e.g., the factor 'performance' was defined by Babar et al. as: *"How performance (productivity/effectiveness) of staff in carrying out the projects help to maintain trusts with clients"*, while in Oza et al. the same term was defined as: *"You have to perform the work to gain the trust, it is based on performance"*. Another example is 'communication', which in both papers is defined as *"How effectiveness of communication with clients [maybe in clients' native language] help to maintain the trusts"*, while we in Oza et al. found three additional terms (transparency, honesty, and understanding) which were used to together represent what Babar et al. referred to as communication.

However, in both cases important terms were well defined, which helped with understanding the differences between them. The Norwegian and Swedish teams were conscious about the definitions. In the cross-case analysis, the results table also includes pairing of the definitions across the two studies. In the thematic synthesis, the definitions were kept in the thematic network so the researchers could always see and compare the different definitions. In the narrative synthesis, the authors were more conservative when aggregating or redefining concepts. Consequently, the narrative analysis concludes that the two studies *"identified entirely different factors"*, while the other two analyses, when analyzing the more detailed meaning of each term, found less differences. Table 8 summarizes the main factors found when using each method of synthesis.

TABLE 8 - COMPARISON OF RESULTS

Thematic Synthesis	Cross-Case Analysis	Narrative Synthesis
Initial Factors	Initial Factors	Establishing Trust

1) Investments in People, Technologies and infrastructure.
  a) Investments
  b) Pilot Projects
  c) Visits
2) Reputation
  a) Experience in Managing Outsourcing projects
  b) References
  c) Reputation
    i) Creditability
    ii) Reputation
	Top list of agreed factors
	1) Personal visits
	2) Investments
	3) People background/experiences
	
	4) Creditability (references, experiences, reputation)
	
	Factors from one study
	1) Cultural understanding
	2) Pilot project performance
	3) Representatives
		Similarities
		1) Client Visits
		2) Investment in infrastructure, processes and human resources
		
		3) Creditability that includes the concepts of:
		  a) Customer References
		  b) Experience in Outsourcing
		  c) Reputation
		
		Differences:
		  Views on Cultural Understanding

Maintaining Trust	Maintaining Factors	Maintaining Trust

1) Defined Development Process
2) Communication
 	
3) Team Performance
4) Commitment
  a. Managing Expectations
  b. Contract Conformance
  c. Confidentiality
  d. Commitment.
	Top list of agreed factors
	1) Development process (consistency)
	2) Communication (honesty, transparency, understanding)
	3) Performance
	4) Demonstrability (quality, timely delivery)
	5) Commitment (managing expectation)
	
	Factors from one study
	1) Cultural understanding
	2) Capabilities
	3) Contract conformance
	4) Personal relationships
	5) Confidentiality
		Similarities:
		1) Processes
		2) Communication
		
		3) Performance
		4) Honesty, Commitment, Confidentiality, Cooperation and Understanding and Contract Conformance, Managing Expectations and Personal Relationship.
		
		Differences:
		  Views on Cultural Understanding

Initial and Maintaining Factors	Cause-effect relations	

1) People Background
  a) Capabilities
  b) People Background
2) Cultural Understanding
  a) Understanding
  b) Cultural Understanding
	1) References, experience, CMM level lead to reputation and creditability, which initiate trust.
	2) Process, project tool, demonstrability and honesty provide transparency that maintains trust.
	3) Communication maintains trust directly, and indirectly via cultural understanding.	

The thematic synthesis method produced a graph showing the relations between the concepts identified, with legends showing which 'trust' factors originate from one study, the other, or both. The cross-case analysis method produced tables, comparing the characteristics of the two cases, and comparing the 'trust' factors originating from the two studies. However, the toolbox of the cross-case analysis method did not enforce aggregation of factors into higher-level factors, as did the thematic synthesis, thus resulting in a longer list of more specific factors. The narrative synthesis produced a text explaining the commonalities and differences in the results from the papers.

Reassuringly, the conclusions on the synthesis of the two papers on factors of trust in outsourcing relationships were largely similar across the thematic and textual narrative synthesis. It is not clear from

our study whether the variations are due the different methods or the two sets of analysts. The conclusions were dominated by the similarities of the results from each paper. The narrative synthesis focused on the differences between the findings in the studies. The thematic synthesis created a new category of factors "initial and maintain" where these differences were placed. One important distinction in the conclusions from the narrative and thematic synthesis was in the interpretation of the factor "Understanding" from the Oza et al. study, which is defined as "Understanding between clients and vendors in transacting with each other". The researchers on the thematic synthesis understood this definition to include the definition of "Cultural Understanding" from Babar et al: "How knowledge of the norms, beliefs, business ethos, and skill in the native language of potential clients helps vendors achieve trusts". Nevertheless, Babar et al. did not consider these definitions as similar. But the analysts on the thematic synthesis concluded that Cultural Understanding was a factor of influence to initial and maintaining trust based on their interpretation of the factor.

The narrative synthesis approach was particularly useful in describing the differences and in making explicit the diversity in the study's context. The thematic synthesis did not extrapolate as much and did not discuss much the contexts of the findings, this method showed rather poor at examining contradictions in the data and at highlighting gaps in the evidence. The cross-case analysis focused more on the semantic similarities and differences between the two studies. The results are hence primarily a synthesized list of factors, expressed in a common language. Further the cause-effect graphing provides an initial understanding of casual relations between factors. However, the lack of access to raw data prevents the analysis from going deeper than the narrative analysis did.

On the transparency of the synthesis, the thematic synthesis and the cross-case analysis showed to fulfill the expectations, but transparency remained a problem in the narrative synthesis, for example, the choice of the examples and quotes in the narrative synthesis rely on the judgments of the researchers, thus it is not clear if they chose the quotes to e.g. reinforce the results from their own studies, or if they chose the quotes that best represented the factors. In the thematic synthesis all the information is traceable and the whole process can be repeated. The same is with the cross-case analysis, where all the matrices and charts can be remade. All the products of the synthesis from the thematic and cross-case analysis can be debatable and discussed again, but that is not the case with the narrative synthesis.

The methods showed to be complementary in some points. For example, the tables of the context on the cross-case analysis counteract the lack of explicit focus on the context of the studies in the thematic synthesis. The thematic synthesis process also led the analysts to extrapolate on the evidence found in each paper and draw conclusions based on the papers. This was not a step foreseen by the cross-case analysis, but it was a natural step conducted by the narrative synthesis.

B. *Challenges of Synthesizing Case Studies*

One important point to highlight is that no matter the method of synthesis, the experience of the analysts will highly influence the final conclusions of the synthesis and that should be accounted for when comparing the methods of synthesis. In addition, by performing the worked example, we could perceive some other factors that can impact the use of a method of synthesis, including: goals and research questions, types and number of the case studies selected, variations in context, limited access to raw data, and quality of the case studies. Most probably, no single method will offer all the required features for the synthesis, so a combination of methods may often be the best approach. In the following subsections we describe the five most important challenges we identified when synthesizing case studies.

1) *Goals and Research Questions*

Several methods have a broad application to a variety of different questions. It is necessary therefore, to select a synthesis method that is applicable to the underlying study aim and question. Typically, a synthesis focuses on a well-defined question and aims to provide an answer by synthesizing the findings

from a relatively narrow range of quality-assessed studies. A fundamental distinction regarding the objective of such syntheses is whether they attempt to provide *knowledge support* or *decision support* [1][26]. A synthesis directed to knowledge support will typically bring together and synthesize evidence on a particular topic, while a synthesis aimed at decision support will be more specific and include analytical tasks to help make a decision within a particular context [18].

Although these are the two ultimate goals, the synthesis goal may vary from the need of pure factual knowledge to attainment of judgment and decision [1], e.g. impacts of objects of study, comparison of objects of study, feasibility of objects of study, impacts of context on the object of study, etc. An object of study can be a technique, a method, an approach, or a phenomenon.

Knowledge of facts, such as whether a specific object of study is important or not, can be suitably answered by a *thematic synthesis* which can bring broad conclusions and is flexible to the buildup of knowledge [1]. Contextualizing an object by comparing different usage contexts can be for example performed with a *cross-case analysis*, as seen in our worked example. Impacts of an object of study on software development as well as the feasibility of the object can be synthesized by thematic synthesis and cross-case analysis. Some more specific techniques and a more interpretative approach would be needed to provide guidelines for decision support; thematic synthesis can be extended for that. In our worked example, it was important that we kept in mind the research question in all the steps of the synthesis, so we would not start exploring other aspects that appeared during the synthesis, as for example exploring how the CMMI context influenced in the answers of the interviews, which would possibly lead to another synthesis process.

2) Number of Case Studies

Some synthesis methods require more studies than others to be effectively applied, e.g. case surveys are tailored to synthesize many studies. However, it is not possible to say for sure how many studies are needed to answer a specific research question. For qualitative studies, the notion of 'saturation' must be taken into account, i.e., judging whether new studies add more knowledge on the research question. The number of studies needed depends on how broad the research question is and how many independent variables and factors affect the results of the object of study. For example, a *narrative synthesis* cannot be meaningfully performed with a large number of cases, as the data volume would be exceedingly large.

In the case that the synthesis comprises many studies, then the synthesis will probably be more quantitative than qualitative. This is so because whenever one attempts to incorporate a large number of cases into a single synthesis, it will be necessary to reduce the evidence to a smaller number of dimensions [11]. Thematic synthesis is a method that is suitable for this scenario. In our worked example, if we would add another paper there were many possible ways, for example, for the coding process, it would just be needed to add references from the new papers to the already existent codes or adding new codes that came with the new studies. Although, it would be more complicated if the new papers would foster new themes or require a reorganization of the already established themes. Cross-case analysis can help to handle with the contexts of the cases but it requires more organization to get the evidence of the papers in tables; the larger the number of papers, the more complex the matrices and tables are.

No matter the method of synthesis, there is always a trade-off between the ability to generalize and the ability to understand fully all the nuances of individual cases. The use of different methods may result in different conclusions. This is a general issue on studies based on qualitative data and is an effect of the richness and lack of precision of such data. A measure to increase the validity of analysis is to maintain a clear chain of evidence from the primary studies to the synthesized evidence, as in our example; we used tools support to manage the traceability of the references to the themes and models. Without this type of support, it is almost impossible to keep the rigor and transparency of the process.

Another method to increase the validity of the synthesized knowledge could be to involve the authors of the primary studies to review the synthesis.

3) Variation in the context of the primary studies

Context is a central concept in empirical software engineering. It is one of the distinctive features of the discipline and it is an indispensable part of software practice. It is likely responsible for one of the most challenging methodological and theoretical problems: study-to-study variation in research findings [10]. The settings in which practice takes place are rarely, if ever, the same. For example, one software organization will have a different environment or be influenced by different environmental factors to that of another software organization. Thus, Dybå et al. discuss the importance of drawing attention to the who, what, when, where and why of a study [10]:

Gerring identifies two possible styles of co-variational evidence in a case study synthesis: temporal and spatial [11]. Spatial variation refers to case studies that were run by different research groups/authors but with similar objectives and instruments of data collection (an example of the 'Who' and 'Where' dimensions). Besides, if different groups perform the studies, one challenge is that they may have different measurement procedures or definitions etc. Temporal variation refers to development over time: if a research group is running a series of case studies successively, the synthesis must consider context variations over time in the studies that may explain the change. Clearly, cases must be similar to each other in whatever respects might affect the causal relationship that the researcher is investigating, or such differences must be controlled for [11]. Uncontrolled heterogeneity means that cases are "apples and oranges", and that one cannot learn anything about underlying causal processes by comparing their histories.

Under circumstances of extreme case-heterogeneity, the researcher may decide that it is better to focus on a single case or a small number of relatively homogeneous cases [11]. Cross-case evidence drawn from a handful of most-similar cases may be more useful than cross-case evidence of many studies, even though the ultimate interest of investment is in a broader population of cases. The issue of population heterogeneity/homogeneity may be understood, therefore, as a trade-off between the number of cases and the number of variables.

The two case studies investigated in this paper were run by different research groups but seen as similar enough in order to make the analysis possible. The Babar et al study was run after the Oza et al. study was published. In this case some of the definitions of the factors in Babar et al. study may have been influenced by Oza et al. study. Another important context variable in this case is the country in which the studies were performed; the Indian software companies were very much influenced by the CMMI model, while the Vietnamese were starting with the process of getting certification level 2 or level 3 on CMMI. Babar et al stated that: '*Vietnamese practitioners believe that being a relatively new player in the software outsourcing arena, they need to quickly build a reputation for being able to develop quality software by following rigorous and systematic processes. Most respondents reported that they have learned from Indian companies that certification is an important mechanism for building creditability and assists in convincing clients to trust in their capabilities*'. These differences were better pointed by the narrative synthesis, because Babar et al. were immersed in the context of their own study, in contrast to the authors of this paper. Details like this may be overlooked when the primary studies do not describe the important factors moderating or influencing the results of the studies. In these cases no method of synthesis can guarantee the validity of the conclusions.

Terminology and definitions may also differ between studies. In some cases they are quite well defined, which helps, but does not solve the problem. Well-established SE terminology may help addressing this challenge, but the area is not mature enough in this way. Therefore the authors of primary studies should be clearer on the definitions of the concepts used in their papers. The challenge

here is that the underlying factors of interest have different meanings in different contexts (conceptual stretching) or the causal relationships are different in different contexts. For example in the example cited in the previous paragraph, Babar et al. found that Creditability was an important factor of trust defined as "*How references, certifications, previous experiences help to gain trust from clients*", but the "certifications" part of the definition is based on what they heard of the experience of the Indians with CMM and CMMI certifications. Further, as mentioned before, Communication had a wider meaning in the Babar et al. study than in the Oza el al. study.

In the Indian context (Oza et al. study) they defined "Reputation" as a factor of trust meaning: "*Vendor's opinion about how certifications from international companies, successful project histories and other previous achievements lead to a good reputation of the company and in turn if becomes useful in achieving trust from the prospective client*", but when they received answers in the context of Indian companies, the answers were based on their own experience with CMM and CMMI certifications. In the case of the Indian respondents only 6 out of 18 mentioned this as an important factor of gaining trust initially, and in the case of the Vietnamese, 11 out of 12 mentioned creditability as an important factor of initial gaining trust from the clients. Clearly the understanding of the importance of the concept for the gaining of trust in the two contexts varies and should be taken into account in the synthesis of the studies.

4) Limited access to raw data

Synthesis of evidence published in journals and conference proceedings involve a key challenge in the limited access to raw data. Authors generally judge differently which level of information is important to be reported in the papers. Some authors, for example, may be very detailed with the descriptions of context information and others may only give some general information on the context of the studies. In shorter articles it may be that certain aspects are not covered due to space limitations, which could affect the analysis. Getting access to or working with other researchers' (probably disclosed) data is not easy. Therefore, using the final report of the results of the primary studies may hinder the synthesis on the details of the evidence being used.

For example in the Babar et al. study, the authors wrote on the evidence about cultural understanding as a factor of trust: "*Respondents described cultural understanding as knowledge of the norms, beliefs, business ethos, and skill in the native language of the potential client. Vietnamese practitioners believe that familiarity with the culture of a client's country and ability to communicate in the native language of that country can help vendors get prospective clients to feel comfortable in starting business initiatives. One of interviewees elaborated on the importance of cultural understanding: 'Ability to communicate in a client's native language and familiarity with his/her culture can provide the biggest advantages or barriers to achieving initial trust.' *"

In this case, Babar et al. used one of the respondents' answers to justify their conclusion that Vietnamese practitioners believe that familiarity with the culture of a client's country and ability to communicate in the native language of that country can help vendors get prospective clients to feel comfortable in starting business initiatives. It seems like they based their conclusion upon reading all the other respondents' answers and found that this respondent's answer was the one that most explicitly showed that they could conclude this way. And in this case we, as analysts in a synthesis, need to trust that their judgment was done in a systematic and impartial way and that this quote really represents the overall view of the evidence.

In the worked example presented here, it was clear that access to the raw data helped on deeper insights in the narrative synthesis, once that they had access to all the data from the study they performed. Nevertheless, if the analysts are not involved in the primary studies included in the synthesis, there is no easy fix to this problem, but it is important that the analysts account for this fact and assure

that the quality of the papers included in the review are good enough to be sure the evidence is credible and complete.

5) Quality Assesment

As the number of primary studies in a synthesis increases, the variation of the quality of the studies also increases. If the studies, which did not meet a certain number of quality criteria were excluded from the synthesis, then little need to be reported regarding the quality of individual studies. But, if these studies were to be included, then there is a need to comment upon the quality of the individual studies as well as the overall strength of evidence when synthesizing the findings [9]. None of the methods discussed in this paper address these questions. So it is much more up to the experience of the researcher to consider the quality of the studies when drawing the conclusions of the synthesis.

Performing quality evaluation of papers is not a straightforward task. However, there have been several suggestions for quality checklists that can be used to evaluate the quality of empirical studies in SE [9]. But, even using these checklists, Kitchenham et al. **Error! Reference source not found.** found that different reviewers perform differently in assessing the quality of the papers. Nevertheless, they also concluded that the reliability obtained by pairs of judges with a round of discussion is generally quite good. Therefore, it is recommended that more than one researcher perform the quality assessment of papers.

In our worked example, the two studies were very similar and their differences in quality did not impact the synthesis. Hence, we did not perform an evaluation of the quality of the papers.

C. Limitations of the study

The main limitation of the study is the number of primary studies. With only two papers we cannot be sure how well our results will generalize. For example, the papers are very simple case studies, consisting of data from interviews only. Besides, the papers are quite similar in terms of the goals and design of the studies; the Babar et al. study was even designed based upon Oza et al. study. This makes the papers easier to synthesize. On the other hand, for the purpose of this paper, we needed cases that would make it possible to demonstrate the usage of the methods of synthesis and how they would be performed in some case studies. Therefore, the papers were purposely chosen to make it relatively easy to perform a synthesis. And, even though they were chosen this way, we still could identify very important challenges during the synthesis.

Other limitations are that we, as a group of researchers, have extensive experience of empirical software engineering, so our results may be better aligned than those that would have been obtained by a random selection of researchers. Nevertheless, this fact also permitted us to be able to be critical to the process of synthesis and to not only reflect on the results of the syntheses of the two papers, or the differences in background among the researchers performing the syntheses.

Another limitation is that the narrative synthesis was not performed for the purpose of the paper; it was performed by the authors of the second primary study. Therefore all the reflections of the method of synthesis were based on our judgments and in what we could perceive as advantages and disadvantages of the method as performed by the authors. Although this might potentially have introduced bias, we felt that it also have enriched our results, once that we could extract the results from an independent synthesis of the two included papers.

V. Conclusions and Future work

While methods for qualitative synthesis have many similarities, there are clear differences in approach between them. In this paper we showed some similarities, differences, and challenges of using three methods of synthesis applied to one example.

The final conclusions of the syntheses reached by the two teams using *thematic synthesis* and *cross-case analysis* were not the same in all aspects, but give different views of the syntheses of the two papers. So the factors derived as the most important factors for trust in outsourcing relationships were sometimes complimentary and sometimes grouped in different perspectives. But overall, the two teams reached similar conclusions. Additionally, Babar et al. included a *narrative synthesis* (which the teams in this study did not read until after their syntheses) focusing on hypothesized differences between the Indian and Vietnamese contexts, which were not part of the original studies.

There are implications for both the conduct of synthesis in secondary studies and for the notion of the differences between methods. With respect to undertaking synthesis, our experience suggests that the process not only depend on the method of synthesis, but also on other factors that can impact the use of a specific method of synthesis, including: goals and research questions, types and number of the case studies selected, variations in context, limited access to raw data, and quality of the case studies. Thus, we recommend that the analysts should be aware of these challenges and try to account for them during the execution of the synthesis. We also recommend that analysts consider using more than one method of synthesis for sake of reliability of the results and conclusions. However, as with the choice of other research methods, personal preferences, educational background, and experience also play an important role in the choice of methods for case study synthesis.

Future work includes increasing the number of papers in the investigation. Our hypothesis is that other challenges will be discovered and some of the challenges described in this paper will impact the final conclusions of a synthesis more than others. We will also explore other methods of synthesis that may be suitable for synthesizing case studies in SE.

REFERENCES

[1] Ashrafian, H., Darsi A., and Athanasiou, T., Evidence Synthesis: Evolving Methodologies to Optimise Patient Care and Enhance Policy Decisions in Evidence Synthesis in Healthcare, Springer-Verlag, London, 2011.

[2] Babar, M.A., Verner. J.M., and Nguyen, P.T., Establishing and maintaining trust in software outsourcing relationships: An empirical investigation. J. Syst. Softw. 80(9): 1438-1449, 2007.

[3] Bethel, E.C, and Bernard, R.M, "Developments and trends in synthesizing diverse forms of evidence: beyond comparisons between distance education and classroom instruction," Distance Education, 31(3): 231-256, 2010.

[4] Cruzes, D.S, and Dybå, T., "Research synthesis in software engineering: A tertiary study", IST, 53(5): 440-455, 2011.

[5] Cruzes, D.S, and Dybå, T., "Recommended Steps for Thematic Synthesis in Software Engineering", ESEM 2011: 275-284.

[6] Cruzes, D.S, and Dybå, T., "Synthesizing evidence in software engineering research", ESEM 2010, 1-10.

[7] Cruzes, D. S, Dybå, T, Runeson, P and Höst, M.: Case Studies Synthesis: Brief Experience and Challenges for the Future. ESEM 2011: 343-346.

[8] Dybå, T., "Contextualizing Empirical Evidence," *IEEE Software*, 30(1): 81-83, 2013.

[9] Dybå, T. and Dingsøyr, T. 2008. Strength of Evidence in Systematic Reviews in Software Engineering, ESEM'2008, pp 178–187.

[10] Dybå, T., Sjøberg, D. I. K. Cruzes, D. S.: What works for whom, where, when, and why?: on the role of context in empirical software engineering. ESEM 2012: 19-28

[11] Gerring, J. Case Study Research: Principles and Practices, New York: Cambridge Univ. Press, 2007.

[12] Khan, S. and VanWynsberghe, R. (2008). Cultivating the Under-Mined: Cross-Case Analysis as Knowledge Mobilization. Forum Qualitative Sozialforschung / Forum: Qualitative Social Research, 9(1), Art. 34, http://nbn-resolving.de/urn:nbn:de:0114-fqs0801348.

[13] Kitchenham, B.A. and Charters, S., Guidelines for performing Systematic Literature Reviews in Software Engineering, Version 2.3, Keele University, EBSE Technical Report, EBSE-2007-01, 2007.

[14] Kitchenham, B.A, Dybå, T, Jørgensen, M. Evidence-based Software Engineering, Proc. ICSE'04, Edinburgh, Scotland, 23–28 May, 2004, pp. 273–281.

[15] Kitchenham, B. A., D. I. K. Sjøberg, T. Dybå, D. Pfahl, P. Brereton, D. Budgen, M. Höst, and P. Runeson (2012). Three Empirical Studies on the Agreement of Reviewers about the Quality of Software Engineering Experiments. Information and Software Technology 54(8), 804–819.

[16] Larsson, R. (1993): Case Survey Methodology: Quantitative Analysis of Patterns across Case Studies, in: The Academy of Management Journal 36 (6), 1515-1546.

[17] Lucas, W. A. (1974): The case survey method: Aggregating case experience. Santa Monica.

[18] Mays, N., Pope, C. and Popay J., Systematically reviewing qualitative and quantitative evidence to inform management and policy-making in the health field, J. Health Serv. Res. Policy, 10: 6–20, 2005.

[19] Miles M.B. and Huberman, A.M. Qualitative Data Analysis: An Expanded Source Book, Sage, 1994.

[20] Miles M.B. and Huberman, A.M. Qualitative Data Analysis: A Sourcebook of New Methods, Sage, 1984.

[21] Newig, J. and O. Fritsch (2009) The case survey method and applications in political science. APSA 2009 Paper. Available at SSRN: ssrn.com/abstract=1451643 (Toronto).

[22] Nguyen, P.T., Babar, M. A. and Verner, J. M.: Critical factors in establishing and maintaining trust in software outsourcing relationships. EASE 2006, http://ewic.bcs.org/content/ConWebDoc/3709

[23] Oza, N., Hall, T., Rainer, A., Grey, S., 2005. Trust in software outsourcing relationships: an empirical investigation of Indian software companies. In: 9th International Conference on Empirical Assessment in Software Engineering (EASE 2005). Keele, UK.

[24] Oza, N., Hall, T., Rainer, A., Grey, S., Trust in software outsourcing relationships: an empirical investigation of Indian software companies. IST, 48: 345–354, 2006.

[25] Popay, J., Roberts, H., Sowden, A., Petticrew, M., Britten, N., Arai, L, Roen, K., and, Rodgers, M.: Developing methods for the narrative synthesis of quantitative and qualitative data in systematic reviews of effects, Centre for Review and Dissemination, ESRC, 2010, http://www.york.ac.uk/inst/crd/projects/narrative_synthesis.htm. Assessed in Dec. 2012.

[26] Pope, C., Mays, N., and Popay, J. Synthesizing Qualitative and Quantitative Health Evidence: A Guide to Methods, Open University Press, 2007.

[27] Ragin, C.C. The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies, Univ. of California Press, 1987.

[28] Runeson, P and Höst, M., "Guidelines for conducting and reporting case study research in software engineering". Empirical Software Engineering, 14(2):131-164, 2009.

[29] Runeson, P., Höst, M., Rainer, A., and Regnell B. Case Study Research in Software Engineering – Guidelines and Examples. Wiley, first edition, ISBN-10: 1118104358, ISBN-13: 978-1118104354, 2012.

[30] Šmite, D., Wohlin, C., Gorschek, T., and Feldt, R. "Empirical evidence in global software engineering: a systematic review", Empirical Software Engineering, 15(1):91-118, 2010.

[31] Thomas, J., and Harden, A., "Methods for the thematic synthesis of qualitative research in systematic reviews", BMC Medical Research Methodology, 8:45, 2008.

[32] Yin R.K. Case study research. Design and methods, 5th edn. London, Sage, 2014.